Service Guarantees for Window Flow Control[1]

R. L. Cruz C. M. Okino
Department of Electrical & Computer Engineering
University of California, San Diego
La Jolla, CA 92093-0407

Abstract

After presenting some new insight into the concept of "service curves," we derive a "service curve guarantee" for a window flow control protocol with cross-traffic characterized by burstiness constraints. Our approach is convenient for studying the end-to-end behavior of hop-by-hop window flow control, and has an interesting relationship with a linear feedback system under the "min-plus" algebra. For affine burstiness constraints on the cross-traffic, we find that a window size proportional to the sum of the burstiness parameters of the cross-traffic and the user bandwidth delay product is sufficient to maximize guaranteed throughput. In addition, we find that buffers need not be as large as window sizes for lossless operation with large propagation delays.

1 System Model

We will consider the two server system depicted in Figure 1, which models a window flow protocol. Traffic from a source is generated according to a function of time R_0, called a rate function, such that $R_0(t)$ is the instantaneous rate at which traffic is being generated at time t. The traffic from the source feeds a buffer, called the first buffer. Traffic departs the first buffer according to the rate function R_1. We assume that the system is empty at time 0. Let $B_1(t)$ denote the amount of traffic held in the first buffer at time t. Thus

$$B_1(t) = \int_0^t R_0(\alpha)d\alpha - \int_0^t R_1(\alpha)d\alpha \ . \tag{1}$$

A server, called the first server, governs the rate R_1 at which traffic departs the first buffer. The server has a transmission capacity of C bits/sec, so that $R_1(t) \leq C$ for all t. In fact, the first server handles other sources of traffic, called "cross-traffic," at time t at rate $I_1(t)$, where $0 \leq I_1(t) \leq C$; thus $R_1(t) \leq C - I_1(t)$.

Traffic departs the first buffer at rate $R_1(t)$ and feeds a "network element", N^f, which serves traffic in a FIFO manner at rate $R^f(t)$. Recalling that the system is empty at time 0, the amount of traffic held in N^f at time t is thus

$$B^f(t) = \int_0^t R_1(\alpha)d\alpha - \int_0^t R^f(\alpha)d\alpha \ . \tag{2}$$

Network element N^f feeds another buffer, called the second buffer. Traffic departs the second buffer according to the rate function $R_2(t)$. The amount of traffic held in the

[1]This work was supported by the National Science Foundation under grants NCR-91-58618 and NCR 94-15684, by the Army Research Office under grant FRI DAAH04-95-1-0248, and by the Air Force Office of Scientific Research under grant F49620-95-1-0538

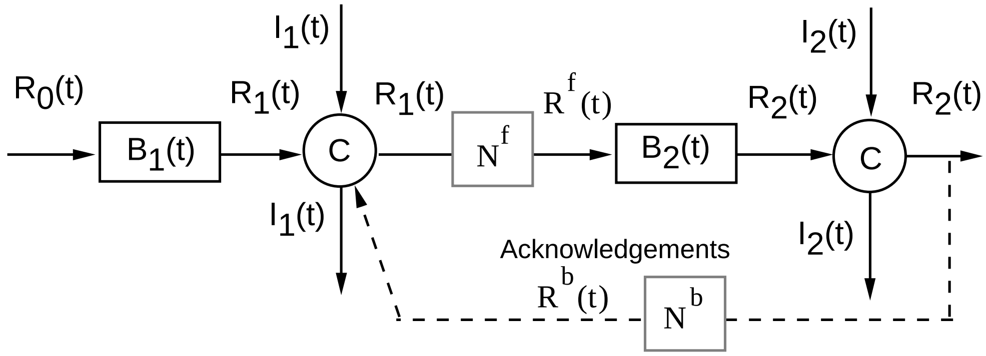

Figure 1: Two Servers in Tandem.

second buffer at time t is thus

$$B_2(t) = \int_0^t R^f(\alpha)d\alpha - \int_0^t R_2(\alpha)d\alpha \ . \tag{3}$$

A second server governs the rate R_2 at which traffic departs from the second buffer. The second server also has a capacity of C bits/sec, and serves cross-traffic from other sources at rate $I_2(t)$ at time t, where $0 \le I_2(t) \le C$; thus $R_2(t) \le C - I_2(t)$.

Traffic departing the first buffer is subject to "window flow control," whereby traffic departing the first buffer may have to wait for acknowledgements from the the second server. More specifically, as traffic from the original source departs the second buffer, acknowledgements are correspondingly generated by the second server and sent back to the first server via a network element N^b. The rate at which acknowledgements are generated at the second server at time t is $R_2(t)$. The network element N^b operates in a FIFO manner and serves acknowledgements at rate $R^b(t)$. Thus, the amount of acknowledgements in N^b at time t is

$$B^b(t) = \int_0^t R_2(\alpha)d\alpha - \int_0^t R^b(\alpha)d\alpha \ . \tag{4}$$

The window flow control protocol operates with respect to a positive parameter K, called the "window size." The first server must insure that no more than K units of traffic are unacknowledged. Thus, traffic in the first buffer may have to wait for acknowledgements to arrive before being eligible for service at the first server. The total amount of traffic unacknowledged (sometimes known as the number of outstanding credits or tokens) at time t is denoted as $T(t)$. Using the definitions above, it follows that

$$\begin{aligned} T(t) &= B^f(t) + B_2(t) + B^b(t) \\ &= \int_0^t R_1(\alpha)d\alpha - \int_0^t R^b(\alpha)d\alpha \ . \end{aligned} \tag{5}$$

The first server serves traffic from the first buffer as fast as possible, but insures that $T(t) \le K$ for all t. More specifically,

$$R_1(t) = \begin{cases} C - I_1(t) & \text{, if } B_1(t) > 0 \text{ and } T(t) < K \\ \min\{C - I_1(t), R_0(t)\} & \text{, if } B_1(t) = 0 \text{ and } T(t) < K \\ \min\{C - I_1(t), R^b(t)\} & \text{, if } B_1(t) > 0 \text{ and } T(t) = K \\ \min\{C - I_1(t), R_0(t), R^b(t)\} & \text{, if } B_1(t) = 0 \text{ and } T(t) = K \ . \end{cases} \tag{6}$$

Similarly, the second server serves traffic from the second buffer as fast as possible:

$$R_2(t) = \begin{cases} C - I_2(t) & \text{, if } B_2(t) > 0 \\ \min\{C - I_2(t), R^f(t)\} & \text{, if } B_2(t) = 0 \ . \end{cases} \tag{7}$$

The system evolution is completely determined by (1)-(7) and the functions R_0, I_1, I_2, R^f, and R^b. We characterize the network elements N^f and N^b by "service curves" S^f and S^b (defined below), respectively. Finally, we assume general "burstiness constraints" [1] on the cross-traffic, i.e. we assume that there exist non-negative non-decreasing functions b_i^{cross} such that for $i = 1, 2$ and all $s \leq t$ there holds

$$\int_s^t I_i(\alpha) d\alpha \leq b_i^{cross}(t - s) \ . \tag{8}$$

1.1 Discussion

The model described above has a number of applications. For example, it may model a single hop within an ATM network that uses credit based flow control. The network element N^f would model forward propagation delay in this case. It has been proposed that acknowledgements only be sent back periodically, rather than continuously, and so the network element N^b could model a combination of backward propagation delay and jitter caused by accumulation of acknowledgements until a burst of acknowledgements is sent back.

Another possible application is the situation in which *end-to-end* window flow control is applied across multiple hops. In this case, N^f would model a combination of queueing and propagation delay suffered in the forward path to the destination. Similarly, N^b would model a combination of queueing and propagation delay in the backward path for acknowledgements returning to the first server.

Alternatively, the model could describe the flow of data across a protocol layer within a host, whereby processes that feed a buffer enter a blocked state when the buffer reaches capacity, and the processes provide bursty service due to multi-tasking within an operating system.

More generally, the model is of interest in manufacturing networks where window flow control is induced naturally by limited storage space for parts being serviced by a sequence of machines, e.g. see [8].

The model described above is a generalization of a model previously proposed in [3], where N^f and N^b were not present (i.e. $R^f \equiv R_1$ and $R^b \equiv R_2$), $b_i^{cross}(x) = \sigma_i + \rho_i x$ (affine burstiness constraints on cross-traffic), and $B_1(0) = \infty$ (infinite supply of packets at the source). For this special case, a lower bound on the throughput of the system, $\liminf_{t\to\infty}(1/t)\int_0^t R_2$, was derived using a "Lyapunov function approach." Although this approach yielded tight results for two servers in tandem, it is difficult to generalize for more than two servers in tandem (see Figure 3) since finding an appropriate Lyapunov function is problematic. Instead, in [3], an aggregation and decomposition technique was used to recursively reduce the analysis of several servers in tandem to the two server case.

In this paper, we use the concept of "service curves" [2] to analyze the system described above. This allows us to easily extend the analysis for more than two servers in tandem, to incorporate "propagation" delays, and to consider general burstiness constraints on the cross-traffic.

Before analyzing the model above, we define service curves and present some new interpretations relating to linear system theory. In Section 3, we present our main results on the two server system discussed in this section. In Section 4, we apply the results of Section 3 to the case of several servers in tandem.

2 Service Curves

Consider a system with entering and exiting traffic described by the rate functions R_{in} and R_{out}. The amount of data stored in the system at time $t \geq 0$ is $B(t) = \int_0^t (R_{in}(\alpha) - R_{out}(\alpha))d\alpha$, where we assume $B(0) = 0$. Suppose S is a given non-negative function. To simplify the notation, assume without loss of generality that $S(x) = 0$ for all $x \leq 0$. Building upon the results in [9], the following definition was essentially proposed in [2] – it is adapted here to the continuous time case.

Definition 1. (Strict Service Curve Guarantee). *A system is said to strictly guarantee the service curve S if for all t, there exists $s \leq t$ with $B(s) = 0$ and $\int_s^t R_{out}(\alpha)d\alpha \geq S(t-s)$.*

The adjective "strict" was not used in [2]. It is added here to differentiate it from a slightly weaker service guarantee which we now introduce:

Definition 2. (Service Curve Guarantee). *A system is said to guarantee the service curve S if for all $t \geq 0$, there exists $s \leq t$ such that $\int_0^t R_{out}(\alpha)d\alpha - \int_0^s R_{in}(\alpha)d\alpha \geq S(t-s)$.*

Note that if $B(s) = 0$, then $\int_0^s R_{in}(\alpha)d\alpha = \int_0^s R_{out}(\alpha)d\alpha$, so that a strict service curve guarantee implies a service curve guarantee. The converse is not necessarily true.

In passing, we note that [10] a (σ, ρ) ("leaky bucket") regulator guarantees the service curve $S_{reg}(x) = \sigma + \rho x$, $x > 0$.

Given two functions F and G defined on the non-negative reals, define the convolution (over the "min-plus algebra") of F and G, written $F * G$, as

$$F * G(x) = \min\{F(x_1) + G(x_2) : x_1 \geq 0, x_2 \geq 0, x_1 + x_2 = x\} ,$$

where the minimum is replaced by an infimum if necessary. It is easy to verify that the convolution operation is commutative and associative, and that it distributes over the minimum operation. It is straightforward to verify that the following is in fact equivalent to Definition 2.

Alternative Definition 2. (Service Curve Guarantee). *A system is said to guarantee the service curve S if for all $t \geq 0$ there holds $\int_0^t R_{out}(\alpha)d\alpha \geq S(t) * \int_0^t R_{in}(\alpha)d\alpha$.*

Define the "impulse function" $\hat{\delta}(x) = 0$ if $x \leq 0$, and $\hat{\delta}(x) = +\infty$ if $x > 0$. Note that for any function F, $F * \hat{\delta}(x) = F(x)$. It is interesting to note then that a service curve is the impulse response of the network element, in some sense, under the min-plus algebra.

In [2], it is shown that if a system strictly guarantees a service curve, then bounds on delay, buffer size, and burstiness of the output traffic are easily derived, assuming burstiness constraints on the traffic arriving to the system. It turns out that identical bounds hold if the service curve guarantee is not necessarily strict. The proofs are almost identical – see [10] for the discrete time case. For completeness, we include these results below. The virtual delay at time t, $D(t)$, is defined as

$$D(t) = \min\{\Delta : \Delta \geq 0 \text{ and } \int_0^{t+\Delta} R_{out}(\alpha)d\alpha \geq \int_0^t R_{in}(\alpha)d\alpha\} .$$

Theorem A. [2] *Assume that* $\int_s^t R_{in}(\alpha)d\alpha \leq b(t-s)$ *for all* $s \leq t$. *Suppose a system guarantees a service curve of* S *(not necessarily strict). Then*
(a) (Buffer Requirements) There holds for all t

$$B(t) \leq \max_{\alpha \geq 0}\{[b(\alpha) - S(\alpha)]^+\} \ .$$

(b) (Bound on Delay) There holds for all t

$$D(t) \leq \max_{\alpha \geq 0}\{\min\{\Delta : \Delta \geq 0 \ \textit{and} \ b(\alpha) \leq S(\alpha + \Delta)\}\} \ .$$

(c) (Output Burstiness) For all $s \leq t$ *there holds* $\int_s^t R_{out}(\alpha)d\alpha \leq b_{out}(t-s)$, *where*

$$b_{out}(x) = \max_{\Delta \geq 0}\{b(x+\Delta) - S(\Delta)\} \ .$$

Theorem B. [2] (Convolution Theorem) *Consider traffic flowing through a system consisting of* n *subsystems in tandem, where the* i^{th} *subsystem guarantees the service curve* S_i *(not necessarily strict). Then the system as a whole guarantees the service curve* $S_{net} = S_1 * S_2 * \cdots * S_n$.

Given the analogy to linear filtering theory, it is natural to ask if there is a concept of a transform domain. Given a function F defined on the non-negative reals, define the "concave conjugate" of F, $\cap^* F$, as

$$\cap^* F(\rho) = \inf_{x : x \geq 0}\{\rho x - F(x)\} \quad , \rho \geq 0 \ .$$

Note that if F is concave, then $\cap^*[\cap^* F] = F$. Similarly, define the "convex conjugate" of F, $\cup^* F$, as

$$\cup^* F(\rho) = \sup_{x : x \geq 0}\{\rho x - F(x)\} \quad , \rho \geq 0 \ ,$$

and note that $\cup^*[\cup^* F] = F$ if F is convex. It is straightforward to verify that for two functions F and G we have

$$\cup^*[F * G](\rho) = \cup^* F(\rho) + \cup^* G(\rho) \ ,$$

which is the analogue of the convolution property for Fourier transforms. It is also interesting to study the mapping from b to b_{out} in Theorem A, part (c). It is easy to verify that the function ρx is an *eigenfunction* of this mapping, in some sense. Furthermore it is not difficult to show that

$$-\cap^* [b_{out}](\rho) \leq -\cap^* [b](\rho) + \cup^*[S](\rho) \ .$$

3 Service Curve for Window Flow Control

We return to the system illustrated in Figure 1. We will assume that network elements N^f and N^b guarantee service curves S^f and S^b respectively (not necessarily strict).

3.1 Service Curve for Window Flow Control

We focus on finding a service curve for the first buffer and server. Given any t, let $s^* = \max\{s : s \leq t \text{ and } B_2(s) = 0\}$. From (7), it follows that $R_2(s) = C - I_2(s)$ for $s \in (s^*, t)$, and therefore from (8) we have

$$\int_{s^*}^{t} R_2(\alpha)d\alpha \geq C(t - s^*) - b_2^{cross}(t - s^*) \ .$$

Since $R_2 \geq 0$, we have

$$\int_{s^*}^{t} R_2(\alpha)d\alpha \geq [C(t - s^*) - b_2^{cross}(t - s^*)]^+ \ .$$

Noting that $B_2(s^*) = 0$, it follows that the subsystem consisting of the second buffer and second server strictly guarantees the service curve S_2, where

$$S_2(x) = [Cx - b_2^{cross}(x)]^+ \ . \tag{9}$$

Note, however, that the subsystem consisting of the first buffer and first server does *not* necessarily guarantee the service curve $\hat{S}_1$ defined as

$$\hat{S}_1(x) = [Cx - b_1^{cross}(x)]^+ \ , \tag{10}$$

since the window flow control protocol may inhibit such a service guarantee. The next theorem identifies a service curve that *is* guaranteed by the subsystem consisting of the first buffer and first server.

Before stating the theorem, we introduce some convenient notation. Given a function G, and a positive integer n, let $G^{(n)}$ be the n-fold convolution of G with itself, i.e.

$$G^{(n)}(x) = \underbrace{G * G * \cdots * G}_{n}(x) \ .$$

In the case where $n = 0$, we define $G^{(0)}(x) = \hat{\delta}(x)$.

Theorem 1 (Service Curve for Window Flow Control) *Suppose that the subsystem consisting of the second buffer and second server guarantees the service curve S_2, and that network elements N^f and N^b guarantee the service curves S^f and S^b, respectively. Define $S_{loop}(x) = \hat{S}_1 * S^f * S_2 * S^b(x)$. Then the subsystem consisting of the first buffer and first server strictly guarantees the service curve S_1, where*

$$S_1(x) = \min_{m \in Z^+} \{\hat{S}_1 * S_{loop}^{(m)}(x) + mK\} \ , \tag{11}$$

and Z^+ is the set of non-negative integers.

> **Remark:** It is interesting to note that that the service curve S_1 given in Theorem 1 is the “impulse response” of the linear feedback system depicted in Figure 2, under the “min-plus” algebra.

To prove Theorem 1 , we will use the lemma below.

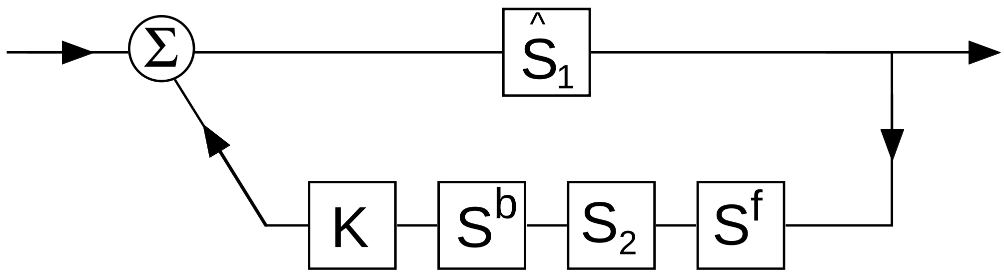

Figure 2: Window Flow Control as a Linear Feedback System.

Lemma 1 *For any fixed t, under the hypothesis of Theorem 1, there exists a finite sequence of intervals $(u_{n+1}, t_n), (t_n, u_n), \cdots, (u_1, t_0)$, with $t_0 = t$, such that $B_1(u_{n+1}) = 0$ and:*

$$\int_{u_{i+1}}^{t_i} R_1(\alpha)d\alpha \geq \hat{S}_1(t_i - u_{i+1}) \quad \forall i = 0, .., n \tag{12}$$

$$\int_{t_{i+1}}^{u_{i+1}} R_1(\alpha)d\alpha \geq S^f * S_2 * S^b(u_{i+1} - t_{i+1}) + K \quad \forall i = 0, .., n-1 \ . \tag{13}$$

Furthermore, the lengths of these intervals satisfy the constraints:

$$t_i - u_{i+1} \geq 0 \ , \quad i = 0, .., n \tag{14}$$

$$u_{i+1} - t_{i+1} \geq K/C \ , \quad i = 0, .., n-1 \ . \tag{15}$$

Proof of Lemma 1: Fix $t \geq 0$ and set $t_0 = t$. We define u_{i+1} in terms of t_i:

$$u_{i+1} = \max\{\max\{s : s \leq t_i, B_1(s) = 0\}, \max\{u : u \leq t_i, T(u) = K\}\} \ . \tag{16}$$

Note that $T(s) < K$ and $B_1(s) > 0$ $\forall s \in (u_{i+1}, t_i)$ and hence using (6) and (8) we have

$$\begin{aligned} \int_{u_{i+1}}^{t_i} R_1(\alpha)d\alpha &= C(t_i - u_{i+1}) - \int_{u_{i+1}}^{t_i} I_1(\alpha)d\alpha \\ &\geq C(t_i - u_{i+1}) - b_1^{cross}(t_i - u_{i+1}) \ . \end{aligned}$$

Remembering that $\int_{u_{i+1}}^{t_i} R_1(\alpha)d\alpha \geq 0$, and using the definition in (10), this implies

$$\int_{u_{i+1}}^{t_i} R_1(\alpha)d\alpha \geq \hat{S}_1(t_i - u_{i+1}) \ ,$$

which proves (12). If $B_1(u_{i+1}) = 0$ then set $n = i$.

Otherwise, if $B_1(u_{i+1}) > 0$, then note that $T(u_{i+1}) = K$. In this case, we will define t_{i+1} in terms of u_{i+1} as follows. By the convolution theorem (Theorem B), the system consisting of the series cascade of N^f, the second buffer and second server, and N^b guarantees a service curve of $S^f * S_2 * S^b(x)$. Thus, there exists $t_{i+1} \leq u_{i+1}$ such that

$$\int_0^{u_{i+1}} R^b(\alpha)d\alpha - \int_0^{t_{i+1}} R_1(\alpha)d\alpha \geq S^f * S_2 * S^b(u_{i+1} - t_{i+1}) \ . \tag{17}$$

Using equations (5) and (17), we get

$$\begin{aligned}
K &= T(u_{i+1}) \\
&= \int_0^{u_{i+1}} R_1(\alpha)d\alpha - \int_0^{u_{i+1}} R^b(\alpha)d\alpha \\
&\leq \int_0^{u_{i+1}} R_1(\alpha)d\alpha - S^f * S_2 * S^b(u_{i+1} - t_{i+1}) - \int_0^{t_{i+1}} R_1(\alpha)d\alpha \ .
\end{aligned} \tag{18}$$

This proves (13).

By construction, we have $t_i - u_{i+1} \geq 0$, which is (14). Since service curves are non-negative and R_1 is bounded above by C, we get from (18)

$$K \leq \int_{t_{i+1}}^{u_{i+1}} R_1(\alpha)d\alpha \leq C(u_{i+1} - t_{i+1}) \ , \tag{19}$$

which implies (15). By construction, $t_{i+1} \geq u_{i+2}$. Hence by adding the non-negative quantity $C(t_{i+1} - u_{i+2})$ to the right side of (19), we get $u_{i+1} - u_{i+2} \geq K/C > 0$. Thus, since the system is completely empty at time 0 and t is finite, the recursion must end in a finite number of iterations, i.e. $B_1(u_{i+1}) = 0$ for some finite i. Thus, the construction results in a finite sequence of intervals as claimed in the lemma.

$\diamondsuit$

Proof of Theorem 1 : Fix any $t > 0$. Invoking Lemma 1, we have

$$\begin{aligned}
\int_{u_{n+1}}^{t} R_1(\alpha)d\alpha &= \sum_{i=0}^{n} \int_{u_{i+1}}^{t_i} R_1(\alpha)d\alpha + \sum_{i=0}^{n-1} \int_{t_{i+1}}^{u_{i+1}} R_1(\alpha)d\alpha \\
&\geq \sum_{i=0}^{n} \hat{S}_1(t_i - u_{i+1}) + \sum_{i=0}^{n-1} \{S^f * S_2 * S^b(u_{i+1} - t_{i+1}) + K\} \\
&= \hat{S}_1(t_n - u_{n+1}) + \sum_{i=0}^{n-1} \{\hat{S}_1(t_i - u_{i+1}) + S^f * S_2 * S^b(u_{i+1} - t_{i+1}) + K\} \\
&\geq \hat{S}_1(t_n - u_{n+1}) + \sum_{i=0}^{n-1} \{S_{loop}(t_i - t_{i+1}) + K\} \\
&\geq \hat{S}_1 * S_{loop}^{(n)}(t - u_{n+1}) + nK \\
&\geq S_1(t - u_{n+1}) \ .
\end{aligned}$$

Noting that $B_1(u_{n+1}) = 0$, this completes the proof.

$\diamondsuit$

Suppose that the delay in network element N^f is upper bounded by τ_f, and that the delay in N^b is upper bounded by τ_b. This would happen if N^f and N^b represent propagation delay. In this case, it is easy to show that for $d = f$ or $d = b$, network element N^d guarantees the service curve S^d, where $S^d(x) = \hat{\delta}(x - \tau_d)$. We use this fact in the next corollary, which considers affine burstiness constraints on the cross-traffic.

Corollary 1 *Suppose $b_1^{cross}(x) = \sigma_a + \rho_a x$, $b_2^{cross}(x) = \sigma_b + \rho_b x$, and the delay through N^f and N^b are bounded by τ_f and τ_b, respectively. Define $\Delta = \tau_f + \tau_b + [\sigma_a/(C - \rho_a)] + [\sigma_b/(C - \rho_b)]$, and $\rho = \max\{\rho_a, \rho_b\}$. Then the system consisting of the first buffer and first server strictly guarantees a service curve of $\bar{S}_1(x)$ where*

$$\bar{S}_1(x) = \min_{m \in Z^+} Q_1^m(x) \ , \tag{20}$$

$$Q_1^0(x) = (C-\rho_a)(x - \frac{\sigma_a}{C-\rho_a})^+ = \hat{S}_1(x) \; , \tag{21}$$

and for $m \geq 1$:

$$Q_1^m(x) = (C-\rho)(x - m\Delta - \frac{\sigma_a}{C-\rho_a})^+ + mK \; . \tag{22}$$

The next corollary assesses the impact of the window size K. Define

$$K^* = \begin{cases} (C-\rho)(\tau_f+\tau_b) + \sigma_a + \frac{C-\rho_a}{C-\rho_b}\sigma_b & \text{, if } \rho_a > \rho_b \\ (C-\rho)(\tau_f+\tau_b) + \frac{C-\rho_b}{C-\rho_a}\sigma_a + \sigma_b & \text{, otherwise.} \end{cases} \tag{23}$$

Corollary 2 *Suppose $b_1^{cross}(x) = \sigma_a + \rho_a x$ and $b_2^{cross}(x) = \sigma_b + \rho_b x$, and the delay through N^f and N^b is upper bounded by τ_f and τ_b respectively. Then the system consisting of the first buffer and first server strictly guarantees a service curve of $\bar{S}_1(x)$, where if $K \geq K^*$ then*

$$\bar{S}_1(x) = \begin{cases} Q_1^0(x) = \hat{S}_1(x) & \text{, if } \rho_a > \rho_b \\ \min\{Q_1^0(x), Q_1^1(x)\} & \text{, otherwise ,} \end{cases} \tag{24}$$

and if $K < K^*$ then

$$\liminf_{x\to\infty} \frac{\bar{S}_1(x)}{x} = \frac{K}{\Delta} \; . \tag{25}$$

> **Remark:** Corollary 2 implies that a window size equal to the sum of the burstiness of the cross-traffic at the first server (σ_a), the burstiness of the cross-traffic at the second server (σ_b), and the user bandwidth delay product ($(C-\rho)(\tau_f+\tau_b)$) is sufficient to guarantee the maximum guaranteed throughput of $C-\rho$.

3.2 Buffer Requirements for Window Flow Control

If the input stream R_0 is characterized by a burstiness constraint, an upper bound on $B_1(t)$ can be derived from Theorem A and Theorem 1. An obvious upper bound for $B_2(t)$ is the window size K, which can be achieved if the delay through network elements N^f and N^b is allowed to be zero. However, it is sometimes possible to derive an upper bound on $B_2(t)$ which is smaller than the window size K, even under no assumptions on R_0.

To illustrate this, we now consider the case where N^f and N^B represent *constant* delays of τ_f and τ_b, respectively. We will derive an upper bound on $B_2(t)$ using two facts. First, note that in this case we have

$$T(t-\tau_f) = \int_0^{t-\tau_f} R_1(\alpha)d\alpha - \int_0^{t-(\tau_f+\tau_b)} R_2(\alpha)d\alpha \;\; \leq \;\; K \; .$$

Thus,

$$\begin{aligned} B_2(t) &= \int_0^{t-\tau_f} R_1(\alpha)d\alpha - \int_0^t R_2(\alpha)d\alpha \\ &\leq \{K + \int_0^{t-(\tau_f+\tau_b)} R_2(\alpha)d\alpha\} - \int_0^t R_2(\alpha)d\alpha \\ &= K - \int_{t-(\tau_f+\tau_b)}^t R_2(\alpha)d\alpha \; . \end{aligned} \tag{26}$$

Second, in this case, the time elapsed from the time a piece of traffic leaves the first server until the acknowledgement for it returns to the first server is at least $\tau_f + \tau_b$. Thus, for any s and t such that $0 \le t - s \le \tau_f + \tau_b$ we have

$$\int_s^t R_1(\alpha)d\alpha \le K \ . \tag{27}$$

Since R_1 is bounded above by the capacity of the server C, from (27) it follows that for any s and t such that $0 \le t - s \le \tau_f + \tau_b$ we have

$$\int_s^t R_1(\alpha)d\alpha \le \min\{C(t-s), K\} \ . \tag{28}$$

Both inequalities (26) and (27) are due to [6].

Theorem 2 *Suppose that N^f and N^b represent constant delays τ_f and τ_b, respectively, and that the cross-traffic I_2 at the second server satisfies the burstiness constraint (8). Then the amount of traffic $B_2(t)$ in the second buffer satisfies*

$$B_2(t) \le \max\{B_2^{\phi_1}, B_2^{\phi_2}\} \ , \tag{29}$$

where

$$\begin{aligned} B_2^{\phi_1} &= \max_{x:0\le x\le \tau_f+\tau_b} \{\min\{Cx, K\} - [Cx - b_2^{cross}(x)]^+\} \\ B_2^{\phi_2} &= K - [C(\tau_f+\tau_b) - b_2^{cross}(\tau_f+\tau_b)]^+ \ . \end{aligned}$$

Proof Fix t and define $u = \max\{s : s \le t, B_2(s) = 0\}$. If $t - u > \tau_f + \tau_b$ then $B_2(s) > 0$ for all $s \in (t - \tau_f - \tau_b, t)$ and hence

$$\begin{aligned} \int_{t-\tau_f-\tau_b}^t R_2(\alpha)d\alpha &= C(\tau_f+\tau_b) - \int_{t-\tau_f-\tau_b}^t I_2(\alpha)d\alpha \\ &\ge [C(\tau_f+\tau_b) - b_2^{cross}(\tau_f+\tau_b)]^+ \ . \end{aligned} \tag{30}$$

Thus, in this case it follows from (26) that $B_2(t) \le B_2^{\phi_2}$.

If $t - u \le \tau_f + \tau_b$ then using (28) it follows that

$$\begin{aligned} B_2(t) &= \int_u^t R^f(\alpha)d\alpha - \int_u^t R_2(\alpha)d\alpha \\ &= \int_{u-\tau_f}^{t-\tau_f} R_1(\alpha)d\alpha - \int_u^t [C - I_2(\alpha)]d\alpha \\ &\le \min\{C(t-u), K\} - [C(t-u) - b_2^{cross}(t-u)]^+ \\ &\le B_2^{\phi_1} \ . \end{aligned}$$

$\diamondsuit$

As an example, suppose that $b_2^{cross}(x) = \sigma_b + \rho_b x$, and $\sigma_b/(C - \rho_b) \le K/C \le \tau_f + \tau_b$. In this case Theorem 2 yields that

$$B_2(t) \le \sigma_b + \rho_b(K/C) \ .$$

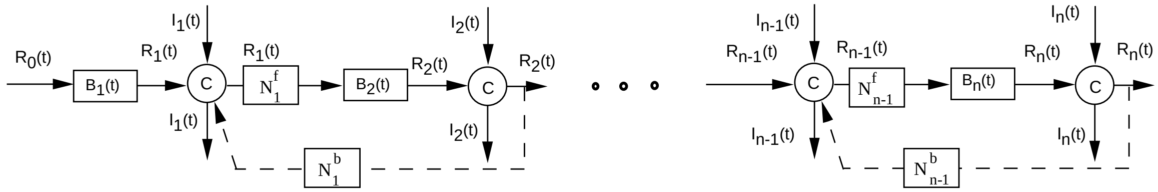

Figure 3: Several Servers in Tandem.

As claimed, the upper bound above may be considerably smaller than the window size K. Smaller upper bounds on $B_2(t)$ can be obtained if R_1 is more strongly regulated. One way to achieve this is to assume a *lower* bound on the cross-traffic at the first server. In this case, if there is insufficient cross-traffic at the first server to achieve this, the first server could "pretend" there is additional cross-traffic. Perhaps a better approach is to have the first server directly control the burstiness of R_1. This is a topic for future research.

4 Several Servers in Tandem

We now investigate the several server case as shown in Figure 3 where we have n cascaded buffer-server pairs. The model is analogous to the two server case and precise definitions are omitted here for brevity.

In order to analyze this multiserver system, we may apply Theorem 1 recursively as follows. Given a burstiness constraint for the cross-traffic at the n^{th} server, a service curve S_n strictly guaranteed by the n^{th} buffer-server pair is implied, analogous to (9). Theorem 1 then implies that a service curve S_{n-1} is strictly guaranteed by the $(n-1)^{th}$ buffer-server pair. This process is repeated until a strict service curve guarantee of S_j is derived for the j^{th} buffer-server pair for all j. The end-to-end service curve for the entire system is then $(S_1 * S_2 * \cdots * S_n) * (S_1^f * S_2^f * \cdots * S_n^f)(x)$ by Theorem B. The following corollary follows by using this method.

Corollary 3 *Consider the tandem configuration of n buffer-server pairs as illustrated in Figure 3, where network elements N_j^f and N_j^b have maximum delays τ_j^f and τ_j^b, respectively. Suppose $b_i^{cross}(x) = \sigma_i + \rho_i x$ for all $i = 1,..,n$, where $\rho = \rho_1 = \rho_2 = ... = \rho_n$.*

If $K_j \geq \sigma_j + \sigma_{j+1} + (C-\rho)(\tau_b^j + \tau_f^j) \quad \forall j = 1,..n-1$, then the system consisting of the j^{th} buffer and the j^{th} server strictly guarantees a service curve of

$$S_j(x) \quad = \quad [-\sigma_j + (C-\rho)x]^+ \quad \forall \quad j = 1, ..., n \ .$$

Furthermore, the entire system guarantees a service curve of $S_{total}(x) = (C-\rho)(x - \sum_{j=1}^{n} \frac{\sigma_j}{C-\rho} - \sum_{j=1}^{n-1} \tau_f^j)^+$. If R_0 is such that $\int_s^t R_0(\alpha)d\alpha \leq \sigma_0 + \rho_0(t-s)$ for all $s \leq t$, where $\rho_0 \leq C-\rho$, then the total end-to-end delay is bounded above by D_{total}, where

$$D_{total} = \frac{\sum_{j=0}^{n} \sigma_j}{C-\rho} + \sum_{j=1}^{n-1} \tau_f^j \ . \tag{31}$$

The results in Corollary 3 improve previously reported bounds, where we showed that $\sum_{j=1}^{n-1} K_j = O(n^2)$ was sufficient for a guaranteed throughput of $C - \rho$ [3] and for a maximum total end-to-end delay bounded by $O(n^2)$ [4].

References

[1] R. Cruz, "A Calculus for Network Delay, Part I: Network Elements in Isolation," *IEEE Transactions on Information Theory*, Vol. 37, No. 1, Jan. 1991.

[2] R. Cruz, "Quality of service guarantees in virtual circuit switched networks," *IEEE Journal on Selected Areas in Communications*, Vol. 13, No. 6, Aug. 1995.

[3] R. Cruz, "Guaranteed throughput for window flow control in a multi-tasking environment," Proceedings of the 1991 Conference on Information Sciences and Systems, John Hopkins University, March 1991.

[4] R. Cruz, 'Tandem queues with finite buffers in a multi-tasking environment: performance guarantees via rate control," Proceedings of the 1992 Conference on Information Sciences and Systems, Princeton University, March 1992.

[5] E. L. Hahne, "Round robin scheduling for fair flow control in data communications networks," *IEEE JSAC*, vol. 9, no. 7, Sept. 1991, pp. 1024-1039.

[6] S. Khorsandi and A. Leon-Garcia, "Robust non-probabilistic bounds for delay and throughput in credit-based flow control," Proceedings IEEE INFOCOM '96, vol. 2, pp 577-84.

[7] H. T. Kung, T. Blackwell, and A. Chapman, "Credit update protocol for flow controlled ATM networks: statistical multiplexing and adaptive credit allocation," Proceedings of SIGCOMM'94, pp. 101-1115.

[8] S. M. Meerkov and F. Top, "Analysis and synthesis of asymptotically reliable serial production lines," Technical Report CGR-24, Dept. of EECS, University of Michigan, Nov. 1989.

[9] A. K. Parekh and R. G. Gallager, "A generalized processor sharring approach to flow control in integrated services networks: the single node case," *IEEE/ACM Trans. Networking,"* vol. 1, no. 3, June 1993, pp. 344-57.

[10] H. Sariowan, "A service-curve approach to performance guarantees in integrated-service networks," Ph.D. thesis, Dept. of Electrical & Computer Engineering, UCSD, June 1996.